Sorting

M269 Tutorial

Contents

1 M269 Tutorial Agenda — Sorting, Recursion

- Welcome & introductions
- *Tutorial topics:* Sorting Algorithms, Recursion
- *Adobe Connect* — if you or I get cut off, wait till we reconnect (or send you an email)
- *Time:* about 1.5 hours
- Do ask questions or raise points.
- *Source:* of slides, notes, programs and playing cards:

  M269Tutorial20250105SortingPrsntn2024J/

  www.pmolyneux.co.uk/OU/M269FolderSync/M269TutorialNotes/M269Tutorial20250105SortingPrsntn2024J/

Introductions — Phil

- *Name* Phil Molyneux
- *Background*
  - Undergraduate: Physics and Maths (Sussex)
  - Postgraduate: Physics (Sussex), Operational Research (Brunel), Computer Science (University College, London)
  - Worked in Operational Research, Business IT, Web technologies, Functional Programming
- *First programming languages* Fortran, BASIC, Pascal
- *Favourite Software*
  - Haskell — pure functional programming language
  - Text editors TextMate, Sublime Text — previously Emacs
  - Word processing in LaTeX — all these slides and notes

  - Mac OS X

- *Learning style* — I read the manual before using the software

Introductions — You

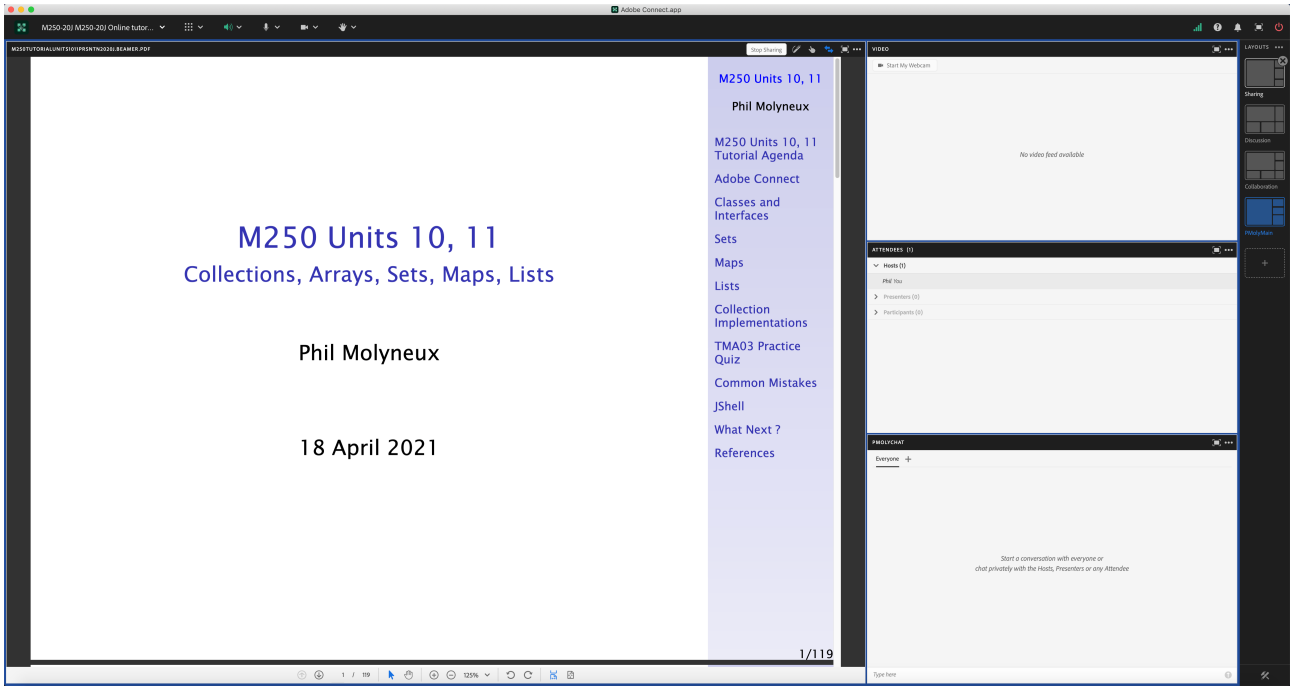
- *Name* ?
- *Favourite software/Programming language* ?
- *Favourite text editor or integrated development environment (IDE)*
- List of text editors, Comparison of text editors and Comparison of integrated development environments
- *Other OU courses* ?
- *Anything else* ?

Go to Table of Contents

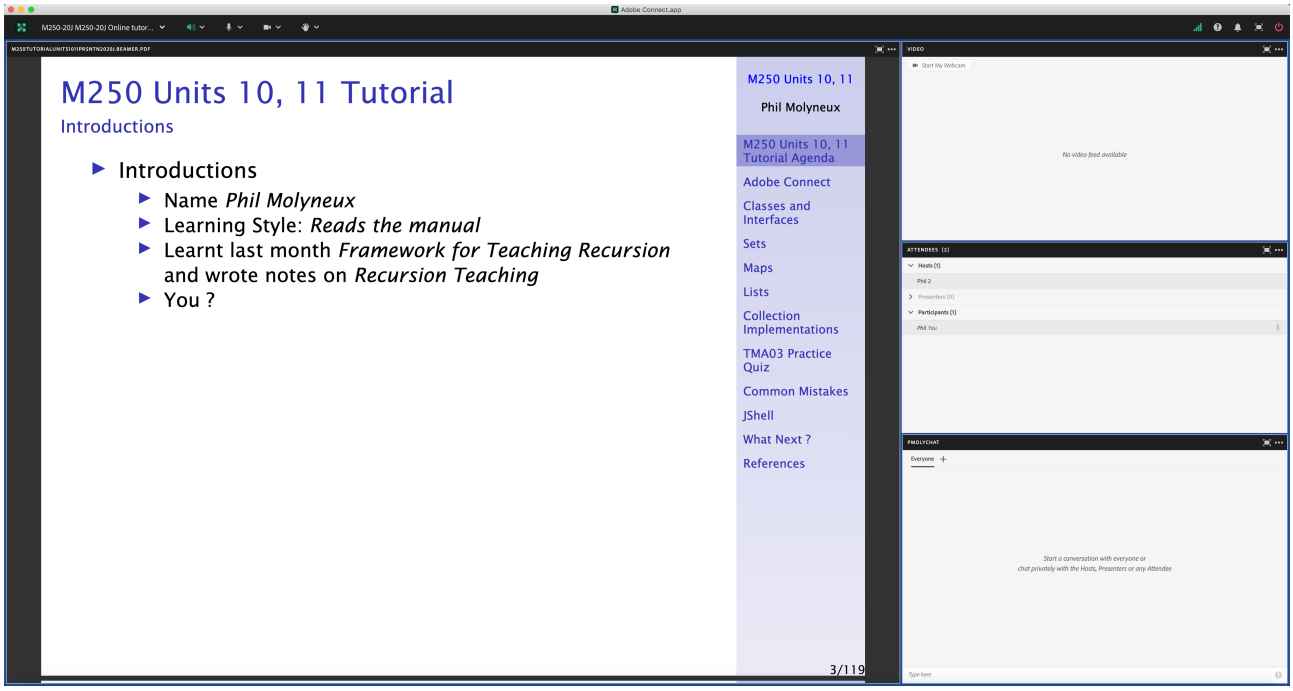
2 Adobe Connect Interface and Settings

2.1 Adobe Connect Interface

Adobe Connect Interface — Host View



Adobe Connect Interface — Participant View



2.2 Adobe Connect Settings

Adobe Connect — Settings

- **Everybody** *Menu bar* › *Meeting* › Speaker & Microphone Setup
- *Menu bar* › *Microphone* › Allow Participants to Use Microphone ✔
- Check Participants see the entire slide including slide numbers bottom right **Workaround**
  - *Disable Draw* *Share pod* › *Menu bar* › *Draw icon*
  - *Fit Width* *Share pod* › *Bottom bar* › *Fit Width icon* ✔
- *Meeting* › Preferences › General › **Host Cursor** › Show to all attendees
- *Menu bar* › *Video* › Enable **Webcam** for Participants ✔
- Do not *Enable single speaker mode*
- Cancel hand tool
- Do not enable green pointer
- **Recording** Meeting › Record Session ✔
- **Documents** Upload PDF with drag and drop to share pod
- Delete *Meeting* › Manage Meeting Information › Uploaded Content and *check filename* › *click on delete*

Adobe Connect — Access

- **Tutor Access**

  TutorHome › M269 Website › Tutorials

  Cluster Tutorials › M269 Online tutorial room

  Tutor Groups › M269 Online tutor group room

Module-wide Tutorials > M269 Online module-wide room

- **Attendance**

  TutorHome > Students > View your tutorial timetables

- **Beamer Slide Scaling** 440% (422 x 563 mm)
- **Clear Everyone's Status**

  Attendee Pod > Menu > Clear Everyone's Status

- **Grant Access** and send link via email

  Meeting > Manage Access & Entry > Invite Participants...

- **Presenter Only Area**

  Meeting > Enable/Disable Presenter Only Area

Adobe Connect — Keystroke Shortcuts

- Keyboard shortcuts in Adobe Connect
- **Toggle Mic** ⌘ + M (Mac), Ctrl + M (Win) (On/Disconnect)
- **Toggle Raise-Hand status** ⌘ + E
- **Close dialog box** ⎋ (Mac), Esc (Win)
- **End meeting** ⌘ + \

2.3 Adobe Connect — Sharing Screen & Applications

- Share My Screen > Application tab > Terminal for Terminal
- Share menu > Change View > Zoom in for mismatch of screen size/resolution (Participants)
- (Presenter) Change to 75% and back to 100% (solves participants with smaller screen image overlap)
- Leave the application on the original display
- Beware blued hatched rectangles — from other (hidden) windows or contextual menus
- Presenter screen pointer affects viewer display — beware of moving the pointer away from the application
- First time: System Preferences > Security & Privacy > Privacy > Accessibility

2.4 Adobe Connect — Ending a Meeting

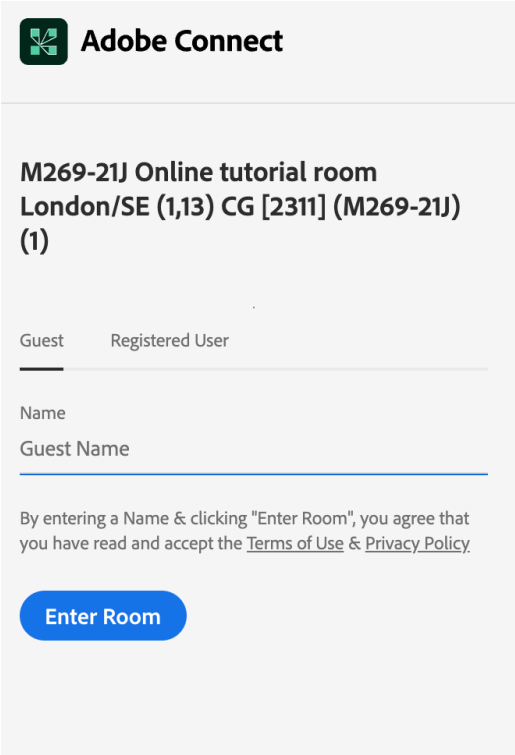
- *Notes for the tutor only*
- **Student:** Meeting > Exit Adobe Connect
- **Tutor:**
- **Recording** Meeting > Stop Recording ✔
- **Remove Participants** Meeting > End Meeting... ✔

  - Dialog box allows for message with default message:
  - *The host has ended this meeting. Thank you for attending.*

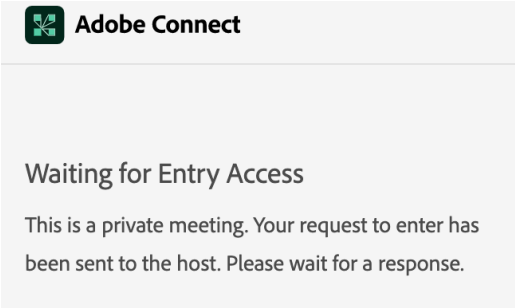
- **Recording availability** *In course Web site for joining the room, click on the eye icon in the list of recordings under your recording* — edit description and name
- **Meeting Information** Meeting › Manage Meeting Information — can access a range of information in Web page.
- **Delete File Upload** Meeting › Manage Meeting Information › Uploaded Content tab select file(s) and click Delete
- **Attendance Report** see course Web site for joining room

2.5 Adobe Connect — Invite Attendees

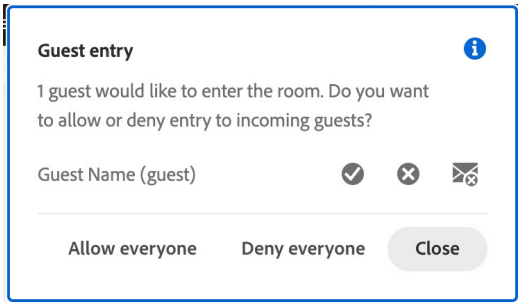
- **Provide Meeting URL** Menu › Meeting › Manage Access & Entry › Invite Participants. . .
- **Allow Access without Dialog** Menu › Meeting › Manage Meeting Information provides new browser window with *Meeting Information* Tab bar › Edit Information
- Check *Anyone who has the URL for the meeting can enter the room*
- Default *Only registered users and accepted guests may enter the room*
- **Reverts to default next session but URL is fixed**
- Guests have blue icon top, registered participants have yellow icon top — same icon if URL is open
- See Start, attend, and manage Adobe Connect meetings and sessions
- Click on the link sent in email from the Host
- Get the following on a Web page
- As *Guest* enter your name and click on Enter Room



- See the *Waiting for Entry Access* for *Host* to give permission



- *Host* sees the following dialog in *Adobe Connect* and grants access



2.6 Layouts

- **Creating new layouts** example *Sharing* layout
- Menu > Layouts > Create New Layout... > Create a New Layout dialog > Create a new blank layout and name it *PMolyMain*
- New layout has no Pods but does have Layouts Bar open (see Layouts menu)
- **Pods**

- Menu › Pods › Share › Add New Share and resize/position — initial name is *Share n* — rename *PMolyShare*
- **Rename Pod** Menu › Pods › Manage Pods... › Manage Pods › Select › Rename or Double-click & rename
- Add Video pod and resize/reposition
- Add Attendance pod and resize/reposition
- Add Chat pod — rename it *PMolyChat* — and resize/reposition
- Dimensions of **Sharing** layout (on 27-inch iMac)
  - Width of Video, Attendees, Chat column 14 cm
  - Height of Video pod 9 cm
  - Height of Attendees pod 12 cm
  - Height of Chat pod 8 cm
- **Duplicating Layouts** does *not* give new instances of the Pods and is probably not a good idea (apart from local use to avoid delay in reloading Pods)
- **Auxiliary Layouts** name *PMolyAux0n*
  - Create new Share pod
  - Use existing Chat pod
  - Use same Video and Attendance pods

2.7 Chat Pods

- **Format Chat text**
- Chat Pod › menu icon › My Chat Color
- Choices: Red, Orange, Green, Brown, Purple, Pink, Blue, Black
- Note: Color reverts to Black if you switch layouts
- Chat Pod › menu icon › Show Timestamps

2.8 Graphics Conversion for Web

- Conversion of the screen snaps for the installation of Anaconda on 1 May 2020
- Using GraphicConverter 11
- File › Convert & Modify › Conversion › Convert
- Select files to convert and destination folder
- Click on Start selected Function or ⌘ + ⏎

2.9 Adobe Connect Recordings

- *Menu bar* › *Meeting* › Preferences › Video
- Aspect ratio › Standard (4:3) (not Wide screen (16:9) default)

- Video quality > Full HD (1080p not High default 480p)
- **Recording** *Menu bar* > *Meeting* > Record Session ✔
- **Export Recording**
- *Menu bar* > *Meeting* > Manage Meeting Information
- *New window* > Recordings > *check Tutorial* > Access Type button
- check Public > check Allow viewers to download
- **Download Recording**
- *New window* > Recordings > *check Tutorial* > Actions > Download File

3 Sorting: Motivation

- Motivation for studying sorting algorithms
- Taxonomy of sorting — see Wikipedia Sorting Algorithm
- Abstract comparison sort — split/join algorithm
- Insertion sort and selection sort described with split/join algorithm diagram and implemented in Python. (A previous edition also included optional Haskell code)
- Recursive and iterative versions
- Mergesort, Quicksort and Bubble sort in the same framework
- Sorting via a data structure — *Tree sort*
- Comparison sorts and Distribution sorts
- Review of Web sites and sorting algorithms used in practice
- From Knuth (1998, page v) *The Art of Computer Programming Vol. 3: Sorting and Searching*
- . . . virtually *every* important aspect of programming arises somewhere in the context of sorting or searching.
- How are good algorithms discovered ?
- How can given algorithms and programs be improved ?
- How can the efficiency of algorithms be analyzed mathematically ?
- How can a person choose rationally between different algorithms for the same task ?
- In what senses can algorithms be proved *best possible* ?
- How does the theory of computing interact with practical considerations ?

3.1 Demonstration 1 Sorting Algorithms as Dances

- AlgoRythmics
- Videos tab > Insertion Sort

- Insertion Sort
- This is the Romanian folk music that inspired Bartók
- Compare the dance with the Python algorithm for Insertion Sort below

ToC

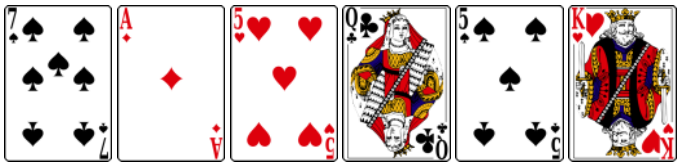
3.2 Activity 1 Card Sorting Exercise

- Almost everyone has played cards and, as part of any card game, will have sorted cards in their hand
- This exercise is aimed at writing down how you sort you cards and giving these instructions to another person to follow.
- Decide on your general ordering of playing cards — you are free to set any ordering you like but here is the usual ordering for suits and values:

```
Clubs < Diamonds < Hearts < Spades

Two < Three < Four < Five < Six
< Seven < Eight < Nine < Ten
< Jack < Queen < King < Ace
```

- Write down your method for sorting cards — the method must specify how to choose a card to move and where to move it to.
- Take the 6 cards given below — record the order of the cards



- Using your method, sort the cards — record the order of the cards after each move of a card
- Now swap your written method and the cards in your original order with another student.
- Follow the other student's method to sort the cards and record your steps

Discussion

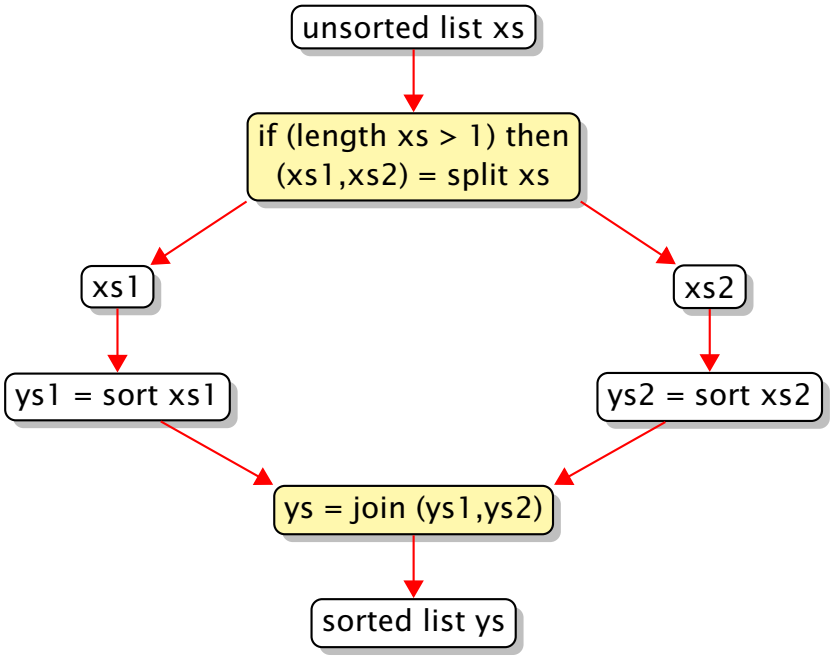
- Did both of you end up with the same sequence of steps?
- Did any of the instructions require human knowledge?
- General point: probably most people use some variation on *Insertion sort* or *Selection sort* but would have steps that had multiple shifts of cards.
- Note: This activity may be done on the *Whiteboard* using cards from http://pmolyneux.co.uk/OU/M269/M269TutorialNotes/M269TutorialSorting/Cards/

ToC

4 Taxonomy of Sorting Algorithms

- Computational complexity — worst, best, average number of comparisons, exchanges and other program contructs (but see http://www.softpanorama.org/Algorithms/sorting.shtml for *Slightly Skeptical View*) — $O(n^2)$ bad, $O(n \log n)$ better
- Other issues: space behaviour, performance on typical data sets, exchanges versus shifts
- Abstract sorting algorithm — Following Merritt (1985); Merritt and Lau (1997) and Azmoodeh (1990, chp 9), we classify the divide and conquer sorting algorithms by easy/hard split/join
- see diagram below

Abstract Sorting Algorithm absSortAlg



4.1 Other Classifications of Sorting Algorithms

- See Wikipedia Sorting algorithm for big list
- *Comparison Sorts*
  - Insertion sort, Selection sort, Merge sort, Quicksort, Bubble sort
  - Sorting via a data structure: Tree sort, Heap sort
- *Non-Comparison sorts* — distribution sorts — bucket sort, radix sort
- *Sorts used in Programming Language Libraries*
  - Timsort by Tim Peters — used in Python and Java — combination of merge and insertion sorts
  - Haskell — modified Mergesort by Ian Lynagh in GHC implementation

ToC

5 Recursion and Iteration

- Many functions are naturally defined using recursion
- A recursive function is defined in terms of calls to itself acting on smaller problem instances along with a base case(s) that terminate the recursion
- Classic example: Factorial $n! = n \times (n-1) \cdots 2 \times 1$

```
5 def fac(n) :
6   if n == 1 :
7     return 1
8   else :
9     return n * fac(n-1)
```

- We can evaluate `fac(6)` by using a substitution model (section 1.1.5) for function application
- *To evaluate a function applied to arguments, evaluate the body of the function with each formal parameter replaced by the corresponding actual arguments.*

  Abelson and Sussman (1996, sec 1.1.5) Structure and Interpretation of Computer Programs

Evaluation of `fac(6)`

	Expression to Evaluate	Reason
	`fac(6)`	Initial line 5
→	`6 * fac(5)`	line 8
→	`6 * (5 * fac(4))`	line 8
→	`6 * (5 * (4 * fac(3))`	line 8
→	`6 * (5 * (4 * (3 * fac(2))))`	line 8
→	`6 * (5 * (4 * (3 * (2 * fac(1)))))`	line 8
→	`6 * (5 * (4 * (3 * (2 * 1))))`	line 6
→	720	Arithmetic

- This occupies more space in the process of evaluation since we cannot do the multiplications until we reach the base case of `fac()`
- This is a *recursive* function and a linear recursive process
- *Implemented* in Python (and most imperative languages) with a *stack* of function calls
- We can define an equivalent *factorial* function that produces a different process

Iterative Factorial

```
24 def facIter(n) :
25   return accProd(n,1)

27 def accProd(n,x) :
28   if n == 1 :
29     return x
30   else :
31     return accProd(n-1, n * x)
```

- `facIter()` uses `accProd()` to maintain a running product and accumulate the final result to return

- We can display the evaluation of `facIter(6)` using the substitution model

Evaluation of `facIter(6)`

	Expression to Evaluate	Reason
	`facIter(6)`	Initial line 24
→	`accProd(6,1)`	line 25
→	`accProd(5, 6 * 1)`	line 30 & (*)
→	`accProd(4, 5 * 6)`	line 30 & (*)
→	`accProd(3, 4 * 30)`	line 30 & (*)
→	`accProd(2, 3 * 120)`	line 30 & (*)
→	`accProd(1, 2 * 360)`	line 30 & (*)
→	720	line 28 & (*)

- This occupies constant space — at each stage all the variables describing the state of the calculation are in the function call
- This is a recursive program and an iterative process
- We are assuming the multiplication is evaluated at each function call (strict or eager evaluation)
- Also referred to as tail recursion — we need not build a stack of calls

Iterative Factorial Exercises

- Write a version of the *factorial* function using a `while` loop in Python
- Write a version of the *factorial* function using a `for` loop in Python

Iterative Factorial Exercises — Solutions

- *Factorial* function using a `while` loop in Python

```
46 def facWhile(n) :
47   x = 1
48
49   while n > 1 :
50     x = n * x
51     n = (n - 1)
52
53   return x
```

- *Factorial* function using a `for` loop in Python

```
57 def facFor(n) :
58   x = 1
59
60   for i in range(n,0,-1) :
61     x = i * x
62
63   return x
```

Tail Recursion and Iteration

- When the structured programming ideas emerged in the 1960s and 1970s the languages such as C and Pascal implemented recursion by always placing the calls on the stack — Python follows this as well

- This means that in those languages they have to have special constructs such as `for` loops, `while` loops, to express iterative processes without recursion
- A `for` loop is syntactically way more complicated than a recursive definition
- Some language implementations (for example, Haskell) spot tail recursion and do not build a stack of calls
- You still have to write your recursion in particular ways to allow the compiler to spot such optimisations.

Structured Programming, GOTO and Recursion

- Böhm and Jacopini (1966) showed that structured programming with a combination of sequence, selection, iteration and procedure calls was Turing complete (see Unit 7)
- In the late 1980s two books came out that were particularly influential:
- Abelson and Sussman (1984, 1996) *Structure and Interpretation of Computer Programs* (known as SICP) which was the programming course for the first year at MIT,
- Bird and Wadler (1988); Bird (1998, 2014) *Introduction to Functional Programming* which was the the programming course for the first year at Oxford.
- See SICP online and Section 1.2 Procedures and the Process They Generate
- Dijkstra (1968) *Go To Statement Considered Harmful* illustrates a debate on structured programming
- The von Neumann computer architecture takes the memory and state view of computation as in Turing m/c
- Lambda calculus is equivalent in computational power to a Turing machine (Turing showed this in 1930s) but efficient implementations did not arrive until 1980s
- Functional programming in Lisp or APL was slow
- Perlis (1982) *Epigrams on Programming*: [Functional programmers] know the value of everything but the cost on nothing
- Meijer et al. (1991) *Recursion is the GOTO of functional programming*
- Leading to common patterns of higher order functions, `map`, `filter`, `fold` and polymorphic data types

ToC

6 Some Split/Join Sorting Algorithms

- Insertion Sort
- Selection Sort
- Merge Sort
- Quicksort
- Bubble Sort

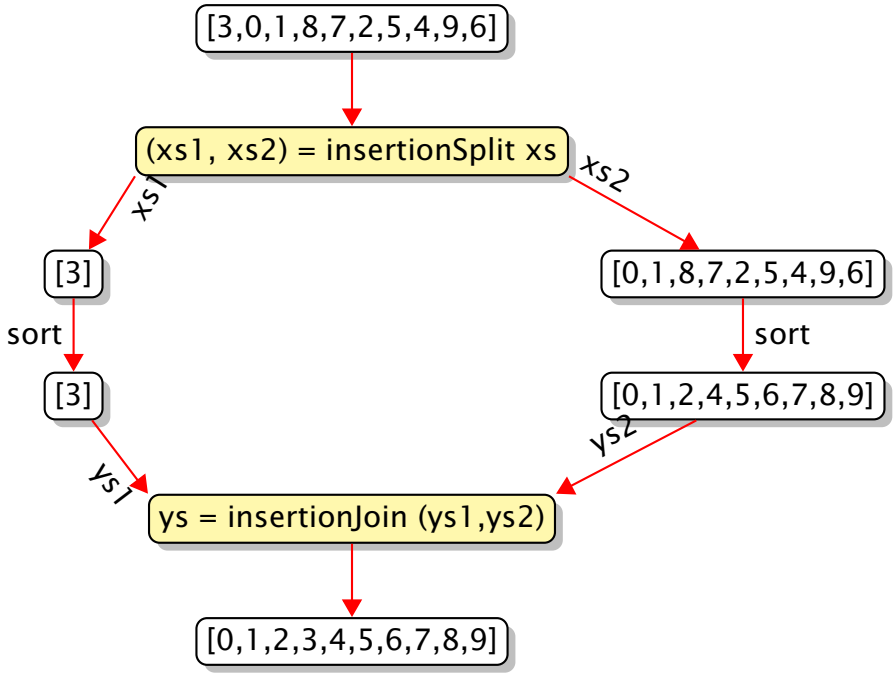
- Implementations in Python, recursive and non-recursive

6.1 Insertion Sort

6.1.1 Insertion Sort — Abstract Algorithm

- *Insertion Split* xs1 is the singleton list of the first item; xs2 is the rest of the list
- *Insertion Join* insert the item in the singleton list into the sorted result of the rest of the list

Abstract Sorting Algorithm Diagram for Insertion Sort insertSortAlg



ToC

6.1.2 Insertion Sort — Python

```
4  def insSort(xs) :
5    if len(xs) <= 1 :
6      return xs
7    else :
8      return ins(xs[0],insSort(xs[1:]))

10 def ins(x,xs) :
11   if xs == [] :
12     return [x]
13   elif x <= xs[0] :
14     return [x] + xs
15   else :
16     return [xs[0]] + ins(x,xs[1:])
```

Python Rewritten

- In the style of the abstract algorithm

```
20 def insSort01(xs) :
21   if len(xs) <= 1 :
22     return xs
23   else :
```

```
24      (xs1,xs2) = insertionSplit(xs)
25      ys1 = insSort01(xs1)
26      ys2 = insSort01(xs2)
27      ys = insertionJoin(ys1,ys2)
28      return ys

30  def insertionSplit(xs) :
31    (xs1,xs2) = (xs[0:1],xs[1:])
32    return  (xs1,xs2)

34  def insertionJoin(ys1,ys2) :
35    if ys2 == [] :
36      return ys1
37    elif ys1[0] <= ys2[0] :
38      return ys1 + ys2
39    else :
40      return ys2[0:1] + insertionJoin(ys1,ys2[1:])
```

ToC

6.1.3 Activity 2 — Insertion Sort: Trace an Evaluation

Insertion Sort — Python Recursive

- Evaluation of `insSort([3,0,1,8,7])`

	Expression to Evaluate	Reason
	`insSort([3,0,1,8,7])`	Initial line 4
→	`ins(3, insSort([0,1,8,7]))`	line 7
→	`ins(3, ins(0, insSort([1,8,7])))`	line 7
→	`ins(3, ins(0, ins(1, insSort([8,7]))))`	line 7
→	`ins(3, ins(0, ins(1, ins(8, insSort([7])))))`	line 7
→	`ins(3, ins(0, ins(1, ins(8, [7]))))`	line 5
→	`ins(3, ins(0, ins(1, ([7] + ins(8, [])))))`	line 15
→	`ins(3, ins(0, ins(1, ([7] + [8]))))`	line 11
→	`ins(3, ins(0, ins(1, [7,8])))`	(+) operator
→	`ins(3, ins(0, ([1] + [7,8])))`	line 13
→	`ins(3, ins(0, [1,7,8]))`	(+) operator
→	`ins(3, ([0] + [1,7,8]))`	line 13
→	`ins(3, [0,1,7,8])`	(+) operator
→	`[0] + (ins 3 [1,7,8])`	line 15
→	`[0] + ([1] + (ins 3 [7,8]))`	line 15
→	`[0] + ([1] + ([3] + ([7,8])))`	line 13
→	`[0,1,3,7,8]`	(+) operator

- Note that the evaluation consumes more space in the process of evaluation;
- also note that you need to be careful with the brackets when doing an evaluation like this by hand.

ToC

6.1.4 Insertion Sort — Non-recursive

- The non-recursive version of *Insertion* sort takes each element in turn and inserts it in the ordered list of elements before it.

```
for index = 1 to (len(xs)-1) do
  insert xs[index] in order in xs[0..index-1]
```

- Here is a Python implementation of the above (based on Miller and Ranum (2011, page 215)).
- It uses the *Python Style Guide PEP 8* http://www.python.org/dev/peps/pep-0008/ (Python Enhancement Proposals) — I must admit I prefer indenting with 2 spaces but I imagine the M269 module will have its own guidelines.

```
42 def insertionSort(xs) :
43   for index in range(1, len(xs)) :
44     currentValue = xs[index]
45     position = index
46     while (position > 0) and xs[position - 1] > currentValue :
47       xs[position] = xs[position - 1]
48       position = position - 1
49
50     xs[position] = currentValue
```

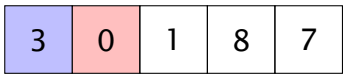
ToC

6.1.5 Activity 3 — Insertion Sort Non-recursive Trace

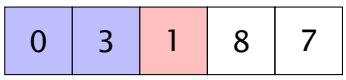
Insertion Sort — Python Non-recursive

- Evaluation of `insertionSort([3,0,1,8,7])`
- Showing just the outer `for index` loop
- 

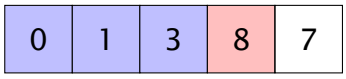
start array



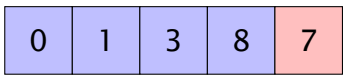
index = 1



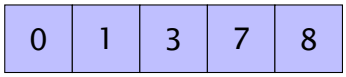
index = 2



index = 3



index = 4



end

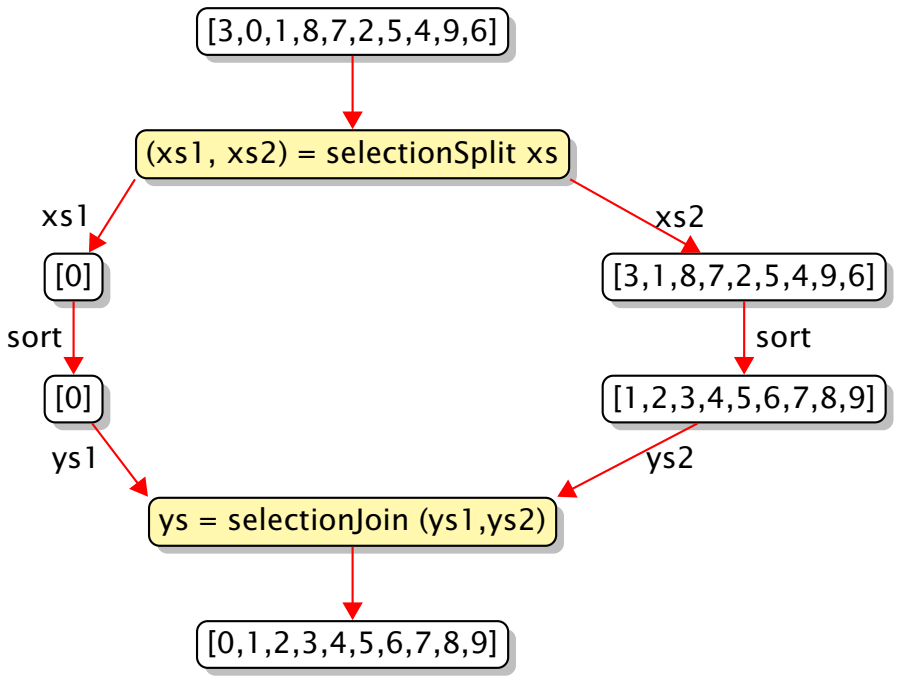
ToC

6.2 Selection Sort

6.2.1 Selection Sort — Abstract Algorithm

- *Selection Split* `xs1` is the singleton list of the minimum item; `xs2` is the original list with the minimum item taken out
- *Selection Join* just put the minimum item and the sorted `xs2` together as the output list

Abstract Sorting Algorithm Diagram for Selection Sort selectSortAlg



ToC

6.2.2 Selection Sort — Python

```
54 def selSort(xs) :
55   if len(xs) <= 1 :
56     return xs
57   else :
58     minElmnt = min(xs)
59     minIndex = xs.index(minElmnt)
60     xsWithoutMin = xs[:minIndex] + xs[minIndex+1:]
61     return [minElmnt] + selSort(xsWithoutMin)
```

- Why do we not use `xs.remove(min(xs))` ?
- *Why do we not use `xs.remove(min(xs))` ?*
- *`remove()` has the side effect of changing the original argument*
- *If we want `selSort()` to be a function, with no side effects then we should use something else*

ToC

6.2.3 Selection Sort — Non-recursive

- The non-recursive version of *Selection* sort takes each position of the list in turn and swaps the element at that position with the minimum element in the rest of the list from that position to the end of the list.

```
for fillSlot = 0 to (len(xs) - 2) do
  find the minimum of
    xs[fillSlot+1]..xs[len(xs) - 1]
  and swap with xs[fillSlot]
```

Selection Sort — Python Non-recursive Implementation

- Here is a Python implementation of the above (based on Miller and Ranum (2011, page 211) but selecting the smallest first not largest, influenced by http://rosettacode.org/wiki/Sorting_algorithms/Selection_sort#PureBasic).
- Note that here we indent by 2 spaces and use the Python idiomatic *simultaneous assignment* to do the swap in line 71

```
63 def selectionSort(xs) :
64   for fillSlot in range(0,len(xs)-1) :
65     minIndex = fillSlot
66     for index in range(fillSlot+1,len(xs)) :
67       if xs[index] < xs[minIndex] :
68         minIndex = index
69
70     # if fillSlot != minIndex: # only swap if different
71     xs[fillSlot],xs[minIndex] = xs[minIndex],xs[fillSlot]
```

M & R Non-recursive Selection SOrt

- The non-recursive version of *Selection* sort in Miller & Ranum sorts in ascending order but takes each position of the list in turn from the right end and swaps the element at that position with the maximum element in the rest of the list from the beginning of the list to that position. (Miller and Ranum, 2011, page 211)

```
  for fillSlot = len(xs) - 1 down to 1 do
    find the maximum of
      xs[0] .. xs[fillSlot]
    and swap with xs[fillSlot]
```

Selection Sort — Python Non-recursive Implementation

- Here is a Python implementation of the above (based on Miller and Ranum (2011, page 211) selecting the largest first.

```
73 def selSortAscByMax(xs) :
74   for fillSlot in range(len(xs) - 1, 0, -1) :
75     maxIndex = 0
76     for index in range(1, fillSlot + 1) :
77       if xs[index] > xs[maxIndex] :
78         maxIndex = index
79
80     temp = xs[fillSlot]
81     xs[fillSlot] = xs[maxIndex]
82     xs[maxIndex] = temp
```

- Note that both Python non-recursive versions work by side-effect on the input list — they do not return new lists.

ToC

6.2.4 Activity 5 — Finding the Non-Recursive Algorithm

- For *Insertion Sort* and *Selection Sort* discuss how the non-recursive case can be found by considering the recursive case and doing the algorithm in place.

ToC

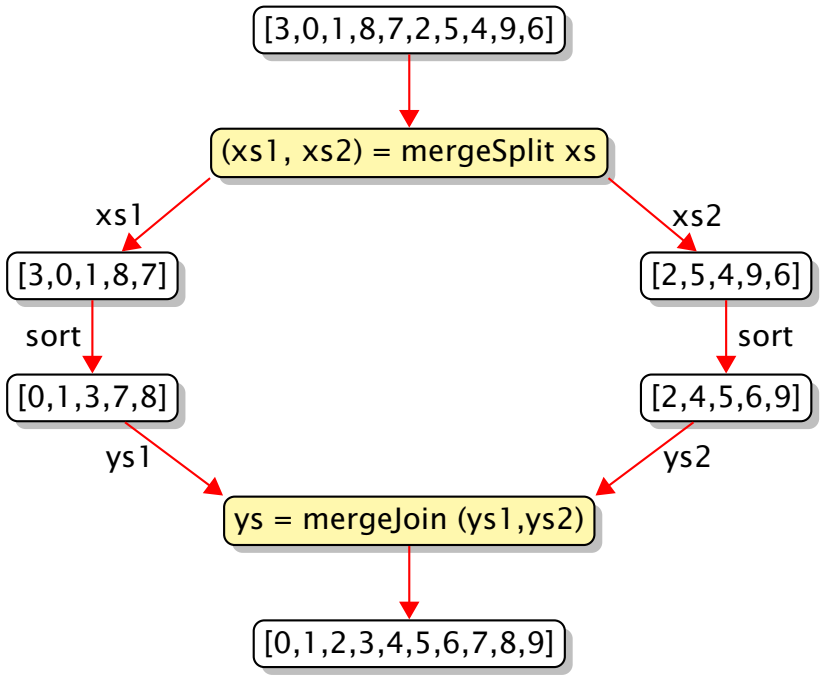
6.3 Merge Sort

6.3.1 Merge Sort — Abstract Algorithm

- *Merge Split* xs1 is half the list; xs2 is the other half of the list.

- *Merge Join* Merge the sorted xs1 and the sorted xs2 together as the output list

Abstract Sorting Algorithm Diagram for Merge Sort mergeSortAlg



ToC

6.3.2 Merge Sort — Python

```
86  def mergeSort(xs) :
87    if len(xs) <= 1 :
88      return xs
89    else :
90      (aList,bList) = mergeSplit(xs)
91      return mergeJoin(mergeSort(aList),mergeSort(bList))

93  def mergeSplit(xs) :
94    return mergeSplit2(xs)

96  def mergeSplit2(xs) :
97    half = len(xs)//2
98    return (xs[:half],xs[half:])

100 def mergeJoin(xs,ys) :
101   if xs == [] :
102     return ys
103   elif ys == [] :
104     return xs
105   elif xs[0] <= ys[0] :
106     return [xs[0]] + mergeJoin(xs[1:],ys)
107   else :
108     return [ys[0]] + mergeJoin(xs,ys[1:])
```

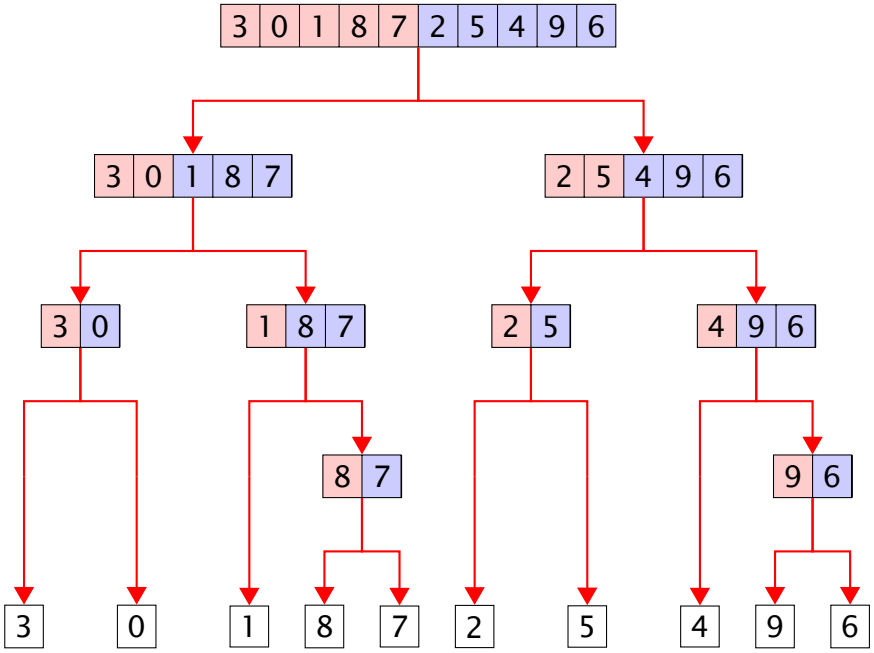
Python `mergeSplit1`

```
110 def mergeSplit1(xs) :
111   if len(xs) == 0 :
112     return  ([],[])
113   elif len(xs) == 1 :
114     return (xs,[])
115   else :
116     (aList,bList) = mergeSplit1(xs[2:])
117     return ([xs[0]] + aList, [xs[1]] + bList)
```

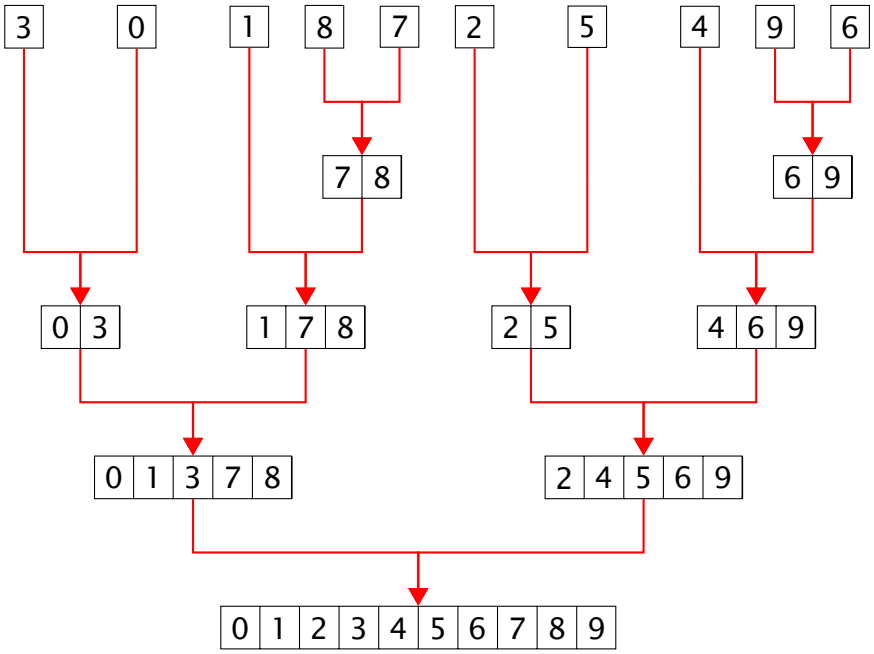
ToC

6.3.3 Merge Sort Diagram

Merge Sort Split Phase mergeSortSplit



Merge Sort Join Phase mergeSortJoin



ToC

6.3.4 Merge Sort Python In-Place

- Here is a Python implementation of the above
- From Miller and Ranum (2011, page 218–221)
- This is also recursive but works in place by changing the array.

- Code from http://interactivepython.org/courselib/static/pythonds/SortSearch/TheMergeSort.html

```
119 def mergeSortInPlace(xs) :
120   if len(xs) > 1 :
121     print("Splitting␣", xs)
122   else :
123     print("Singleton␣", xs)
124
125   if len(xs) > 1 :
126     half = len(xs)//2
127     (aList, bList) = (xs[:half],xs[half:])
128
129     mergeSortInPlace(aList)
130     mergeSortInPlace(bList)
```

```
132     i,j,k = 0,0,0
133     while i < len(aList) and j < len(bList) :
134       if aList[i] < bList[j] :
135         xs[k] = aList[i]
136         i = i + 1
137       else :
138         xs[k] = bList[j]
139         j = j + 1
140       k = k + 1
141
142     while i < len(aList) :
143       xs[k] = aList[i]
144       i = i + 1
145       k = k + 1
146
147     while j < len(bList) :
148       xs[k] = bList[j]
149       j = j + 1
150       k = k + 1
```

- Here is the code that reports the merging of the lists

```
152   if len(xs) > 1 :
153     print("Merging␣", aList, ",", bList, "to", xs)
154   else :
155     print("Merged␣", xs)
```

- ␣ is how the `listings` package shows spaces in strings by default (read the manual)
- `//` is the Python integer division operator
- `aList[start:stop:step]` is a *slice* of a list — see Python Sequence Types — slice operations return a new list (van Rossum and Drake, 2011a, page 19) so `xs[:]` returns a copy (or clone) of `xs` — if any of the indices are missing or negative than you have to think a bit (or read the manual)
- In Python you really do need to be aware when you are working with values or references to objects.
- A listing of the output of `mergeSortInPlace(xsc)` below is given in the article version of these notes

```
>>> from SortingPython import *
>>> xs = [3,0,1,8,7,2,5,4,9,6]
>>> xsc = xs[:]
>>> mergeSortInPlace(xsc)
Splitting  [3, 0, 1, 8, 7, 2, 5, 4, 9, 6]
#
# lines removed
#
Merging  [0, 1, 3, 7, 8] , [2, 4, 5, 6, 9]
  to [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
```

Complete listing of output of mergeSortInPlace(xsc)

```
Python3>>> from SortingPython import *
Python3>>> xs = [3,0,1,8,7,2,5,4,9,6]
Python3>>> xsc = xs[:]
Python3>>> mergeSortInPlace(xsc)
Splitting  [3, 0, 1, 8, 7, 2, 5, 4, 9, 6]
Splitting  [3, 0, 1, 8, 7]
Splitting  [3, 0]
Singleton  [3]
Merged  [3]
Singleton  [0]
Merged  [0]
Merging  [3] , [0] to [0, 3]
Splitting  [1, 8, 7]
Singleton  [1]
Merged  [1]
Splitting  [8, 7]
Singleton  [8]
Merged  [8]
Singleton  [7]
Merged  [7]
Merging  [8] , [7] to [7, 8]
Merging  [1] , [7, 8] to [1, 7, 8]
Merging  [0, 3] , [1, 7, 8] to [0, 1, 3, 7, 8]
Splitting  [2, 5, 4, 9, 6]
Splitting  [2, 5]
Singleton  [2]
Merged  [2]
Singleton  [5]
Merged  [5]
Merging  [2] , [5] to [2, 5]
Splitting  [4, 9, 6]
Singleton  [4]
Merged  [4]
Splitting  [9, 6]
Singleton  [9]
Merged  [9]
Singleton  [6]
Merged  [6]
Merging  [9] , [6] to [6, 9]
Merging  [4] , [6, 9] to [4, 6, 9]
Merging  [2, 5] , [4, 6, 9] to [2, 4, 5, 6, 9]
Merging  [0, 1, 3, 7, 8] , [2, 4, 5, 6, 9] to [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
```

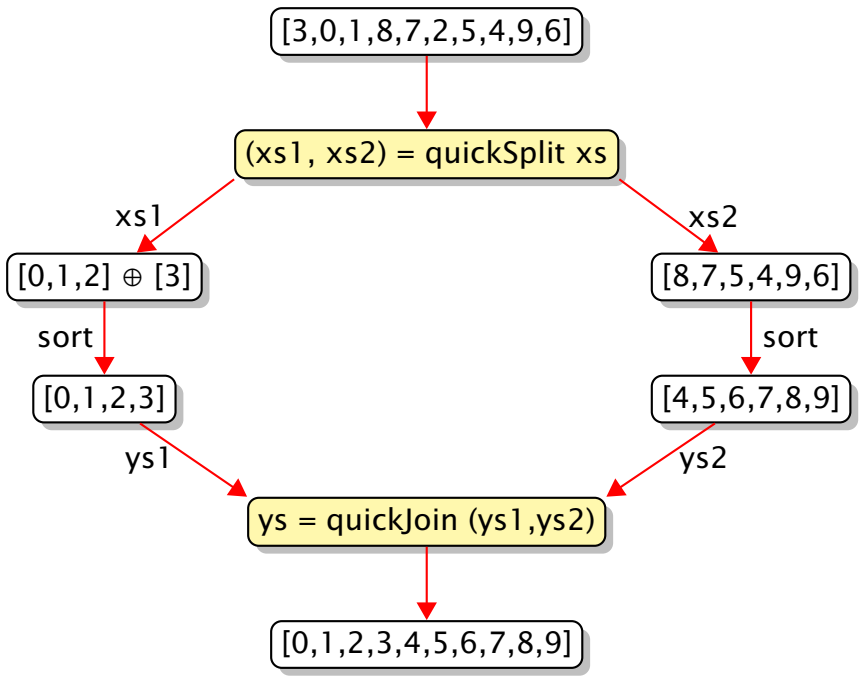
ToC

6.4 Quicksort

6.4.1 Quicksort — Abstract Algorithm

- *Quicksort Split* Choose an item in the list to be the *pivot* item; xs1 comprises items in the list less than the pivot plus the pivot; xs2 comprises items in the list greater than or equal to the pivot.
- *Quicksort Join* just append the sorted xs1 and the sorted xs2 together as the output list

Abstract Sorting Algorithm Diagram for Quicksort quicksortSortAlg



Note: the diagram use ⊕ as the *list append* operator — this is used in various courses and texts

ToC

6.4.2 List Comprehensions

- Haskell 2010 Language Report section 3.11 *List Comprehensions*
- $[e \mid q_1, \ldots, q_n]$, $n \geq 1$ where q_i qualifiers are either
  - *generators* of the form `p <- e` where p is a pattern of type t and e is an expression of type $[t]$
  - *local bindings* that provide new definitions for use in the generated expression e or subsequent boolean guards and generators
  - *boolean guards* which are expressions of type Bool
- Python Language Reference section 6.2.4 *Displays for lists, sets and dictionaries* and section 6.2.5 *List displays*
- `[ expr for target in list]` — simple comprehension
- `[ expr for target in list if condition]` — filters
- `[ expr for target1 in list1 for target2 in list2]` — multiple generators

ToC

6.4.3 Quicksort — Python

```
159 def qsort(xs) :
160   if not xs :
161     return []
162   else :
163     pivot = xs[0]
164     less = [x for x in xs     if x <  pivot]
165     more = [x for x in xs[1:] if x >= pivot]
166     return qsort(less) + [pivot] + qsort(more)
```

- The `if` test at line 160 shows that Python is weakly typed (and the author of this code comes from JavaScript)

ToC

6.4.4 Quicksort Python In-Place

- The in-place version of Quick sort works by partitioning a list in place about a value `pivotvalue`: (Azmoodeh, 1990, page 259–266)

(1) Scan from the left until

   - `alist[leftmark] >= pivotvalue`

(2) Scan from the right until

   - `alist[rightmark] < pivotvalue`

(3) Swap `alist[leftmark]` and `alist[rightmark]`

(4) Repeat (1) to (3) until scans meet

- Here is an in place version of Quick Sort from Miller and Ranum (2011, pages 221–226)
- Code based on http://interactivepython.org/courselib/static/pythonds/SortSearch/TheQuickSort.html

```
168 def quickSort(xs) :
169   quickSortHelper(xs, 0, len(xs) - 1)

171 def quickSortHelper(xs, fst, lst) :
172   if fst < lst :

174     splitPoint = partition(xs,fst,lst)

176     quickSortHelper(xs, fst, splitPoint - 1)
177     quickSortHelper(xs, splitPoint + 1, lst)
```

```
179 def partition(xs,fst,lst) :
180   pivotValue = xs[fst]
181   leftMk     = fst + 1
182   rightMk    = lst
183   done       = False

185   while not done :
186     while leftMk <= rightMk and \
187             xs[leftMk] <= pivotValue :
188       leftMk = leftMk + 1
189     while xs[rightMk] >= pivotValue and \
190             rightMk >= leftMk :
191       rightMk = rightMk - 1

193     if rightMk < leftMk :
194       done = True
195     else :
196       xs[leftMk], xs[rightMk] = xs[rightMk], xs[leftMk]

198   xs[fst], xs[rightMk] = xs[rightMk], xs[fst]
199   return rightMk
```

- The (\) is enabling a statement to span multiple lines — see Lutz (2009, page 317), Lutz (2013, page 378)
- for a language that uses the *offside rule* why do we need to do this?

- Note that using (\) to create continuations is *frowned on* (Lutz, 2009, page 318), Lutz (2013, page 379)
- the authors should have put the entire boolean expression inside parentheses () so that we get implicit continuation.
- This is not mentioned explicitly in the *Style Guide for Python Code* http://www.python.org/dev/peps/pep-0008/ but it does explicitly mention using Python's implicit line joining with layout guidelines.

ToC

6.5 Bubble Sort

6.5.1 Bubble Sort — Abstract Algorithm

- *Bubble sort* is rather like the *Hello World* program of sorting algorithms — we have to include it even it isn't very useful in practice.
- It can be thought of as an in-place version of Selection sort
- In the implementations below, in each pass through the list, the next highest item is moved (bubbled) to its proper place.
- OK, I should have written it to bubble the smallest the other way to be consistent with the implementations of Selection sort above.

ToC

6.5.2 Bubble Sort — Python

- Here is a Python implementation from Miller and Ranum (2011, pages 207–210)
- it does not test if there have been no swaps but does use some knowledge of the algorithm by reducing the pass length by one each time (which the Haskell one did not do)

```
203 def bubbleSort(xs) :
204   for passNum in range(len(xs) - 1, 0, -1) :
205     for i in range(passNum) :
206       if xs[i] > xs[i+1] :
207         xs[i], xs[i+1] = xs[i+1], xs[i]
```

- Note that range() is a built-in function to Python that is used a lot
- Read the documentation at Section 4.6.6 Ranges
- Remember that range(5) means [0,1,2,3,4] (not [0,1,2,3,4,5] or [1,2,3,4,5])

ToC

7 Future Work

Programming, Debugging, Psychology

Although programming techniques have improved immensely since the early days, the process of finding and correcting errors in programming — known graphically if inelegantly as *debugging* — still remains a most difficult, confused and unsatisfactory opera-

tion. The chief impact of this state of affairs is psychological. Although we are happy to pay lip-service to the adage that to err is human, most of us like to make a small private reservation about our own performance on special occasions when we really try. It is somewhat deflating to be shown publicly and incontrovertibly by a machine that even when we do try, we in fact make just as many mistakes as other people. If your pride cannot recover from this blow, you will never make a programmer.

Christopher Strachey, Scientific American 1966 vol 215 (3) September pp112–124

- To err is human, to really foul things up requires a computer.
- Attributed to Paul R. Ehrlich in 101 Great Programming Quotes
- Attributed to Bill Vaughn in Quote Investigator
- Derived from Alexander Pope (1711, An Essay on Criticism)
- *To Err is Humane; to Forgive, Divine*
- This also contains

  A little learning is a dangerous thing;

  Drink deep, or taste not the Pierian Spring
- In programming, this means you have to *read the fabulous manual* (RTFM)

Sorting, Searching, Binary Trees

- Recursive function definitions
- Inductive data type definitions
  - A *list* is either an empty list or a first item followed by the rest of the list
  - A *binary tree* is either an empty tree or a node with an item and two sub-trees
- Recursive definitions often easier to find than iterative
- Sorting
- Searching
- Both use binary tree structure — either implicitly or explicitly

Future Work — Dates

- Sunday 5 January 2025 Tutorial Online Sorting, Recursion
- Sunday 12 January 2025 Tutorial Online Binary Trees, Recursion
- Sunday 9 February 2025 (Module wide) Tutorial Online Binary Trees, Recursion
- Thursday, 13 March 2025 TMA02
- Sunday, 9 March 2025 Tutorial (Online): Graphs, Greedy Algorithms
- Sunday, 6 April 2025 Tutorial (Online): (Module wide) Dynamic Programming
- Sunday, 27 April 2025 Tutorial (Online): (Module wide) Computability, Complexity
- Sunday, 4 May 2025 Tutorial (Online): Review of course material for TMA03

- Thursday, 8 May 2025 TMA03

ToC

8 Web Sites & References

8.1 Sorting Web Links

- **Rosetta Code Sorting Algorithms** http://rosettacode.org/wiki/Sorting_algorithms — sorting algorithms implemented n lots of programming languages
- **Sorting Algorithm Animations** https://www.toptal.com/developers/sorting-algorithms — visual display of the performance of various sorting algorithms for several classes of data: random, nearly sorted, reversed, few unique — worth browsing to.
- **Sorting Algorithms as Dances** https://www.youtube.com/user/AlgoRythmics — inspired!

ToC

8.2 Python Web Links & References

- **Miller and Ranum (2011)** http://interactivepython.org/courselib/static/pythonds/index.html — the entire book online with a nice way of running the code.
- **Lutz (2013)** — one of the best introductory books
- **Lutz (2011)** — a more advanced book — earlier editions of these books are still relevant — you can also obtain electronic versions from the O'Reilly Web site http://oreilly.com
- **Python 3 Documentation** https://docs.python.org/3/
- **Python Style Guide PEP 8** https://www.python.org/dev/peps/pep-0008/ (Python Enhancement Proposals)

ToC

8.3 Haskell Web Links & References

- **Haskell Language** https://www.haskell.org
- **HaskellWiki** https://wiki.haskell.org/Haskell
- **Learn You a Haskell for Great Good!** — very readable introduction to Haskell
- **Real World Haskell** http://book.realworldhaskell.org — more advanced
- **Thompson (2011)** — a good text for functional programming for beginners
- **Bird and Wadler (1988); Bird (1998, 2014)** — one of the best introductions but tough in parts, requires some mathematical maturity — the three books are in effect different editions

- **Bird and Gibbons (2020)** — the descriptions of five main principles of algorithm design: divide and conquer, greedy algorithms, thinning, dynamic programming, and exhaustive search, are mainly language neutral
- **Functors, Applicatives, and Monads in Pictures** — a very good outline with cartoons
- **Typeclassopedia** https://wiki.haskell.org/Typeclassopedia — a more formal introduction to Functors, Applicatives and Monads
- **Haskell Wikibook**

ToC

8.4 Demonstration 2 Sorting Algorithms as Dances

- Quicksort
- https://www.youtube.com/user/AlgoRythmics
- the hats make the point(!)

ToC

References

Abelson, Harold and Gerald Jay Sussman (1984). *Structure and Interpretation of Computer Programs*. MIT Press, first edition. URL http://mitpress.mit.edu/sicp/. 14

Abelson, Harold and Gerald Jay Sussman (1996). *Structure and Interpretation of Computer Programs*. MIT Press, second edition. ISBN 0262510871. URL http://mitpress.mit.edu/sicp/. 12, 14

Azmoodeh, Manoochehr (1990). *Abstract Data Types and Algorithms*. Palgrave Macmillan, second edition. ISBN 0333512103. 11, 25

Bird, Richard (1998). *Introduction to Functional Programming using Haskell*. Prentice Hall, second edition. ISBN 0134843460. 14, 28

Bird, Richard (2014). *Thinking Functionally with Haskell*. Cambridge University Press. ISBN 1107452643. URL https://www.cs.ox.ac.uk/publications/books/functional/. 14, 28

Bird, Richard and Jeremy Gibbons (2020). *Algorithm Design with Haskell*. Cambridge University Press. ISBN 9781108869041. URL https://www.cs.ox.ac.uk/publications/books/adwh/. 29

Bird, Richard and Phil Wadler (1988). *Introduction to Functional Programming*. Prentice Hall, first edition. ISBN 0134841972. 14, 28

Böhm, Corrado and Giuseppe Jacopini (1966). Flow diagrams, Turing Machines and Languages with Only Two Formation Rules. *Communications of the ACM*, 9(5):366–371. 14

Dijkstra, Edsger W (1968). Letters to the editor: Go To Statement Considered Harmful. *Communications of the ACM*, 11(3):147–148. 14

Dromey, R.Geoff (1982). *How to Solve it by Computer*. Prentice-Hall. ISBN 0134340019.

Dromey, R.Geoff (1989). *Program Derivation: The Development of Programs from Specifications*. Addison Wesley. ISBN 0201416247.

Hudak, Paul; John Hughes; Simon Peyton Jones; and Phil Wadler (2007). A History of Haskell: Being Lazy with Class. In *Proceedings of the third ACM SIGPLAN conference on History of programming languages*, pages 12–1–12–55. ACM New York, NY, USA.

Knuth, D.E. (1998). *The Art of Computer Programming Vol. 3: Sorting and Searching*. The Art of Computer Programming: Sorting and Searching. Adddison Wesley, second edition. ISBN 0201896850. URL http://books.google.co.uk/books?id=sXa_mwEACAAJ. 9

Lee, Gias Kay (2013). Functional Programming in 5 Minutes. Web. http://gsklee.im, URL http://slid.es/gsklee/functional-programming-in-5-minutes.

Lutz, Mark (2009). *Learning Python*. O'Reilly, fourth edition. ISBN 0596158068. 25, 26

Lutz, Mark (2011). *Programming Python*. O'Reilly, fourth edition. ISBN 0596158106. URL http://learning-python.com/books/about-pp4e.html. 28

Lutz, Mark (2013). *Learning Python*. O'Reilly, fifth edition. ISBN 1449355730. URL http://learning-python.com/books/about-lp5e.html. 25, 26, 28

Marlow, Simon and Simon Peyton Jones (2010). Haskell Language and Library Specification. Web. URL http://www.haskell.org/haskellwiki/Language_and_library_specification.

Meijer, Erik; Maarten Fokkinga; and Ross Paterson (1991). Functional programming with bananas, lenses, envelopes and barbed wire. In *Functional Programming Languages and Computer Architecture*, pages 124–144. Springer. 14

Merritt, SM and KK Lau (1997). A logical inverted taxonomy of sorting algorithms. In *Proceedings of the Twelfth International Symposium on Computer and Information Sciences*, pages 576–583. Citeseer. 11

Merritt, Susan M (1985). An inverted taxonomy of sorting algorithms. *Communications of the ACM*, 28(1):96–99. 11

Miller, Bradley W. and David L. Ranum (2011). *Problem Solving with Algorithms and Data Structures Using Python*. Franklin, Beedle Associates Inc, second edition. ISBN 1590282574. URL http://interactivepython.org/courselib/static/pythonds/index.html. 17, 19, 21, 25, 26, 28

Perlis, Alan J. (1982). Epigrams on Programming. *SIGPLAN Notices*, 17(9):7–13. 14

Sussman, Julie (1985a). *Instructor's Manual to Accompany Structure and Interpretation of Computer Programs*. MIT Press. ISBN 0262 691019. URL http://mitpress.mit.edu/sites/default/files/sicp/index.html.

Sussman, Julie (1985b). *Instructor's Manual to Accompany Structure and Interpretation of Computer Programs*. MIT Press, second edition. ISBN 0262 692201. URL http://mitpress.mit.edu/sites/default/files/sicp/index.html.

Thompson, Simon (2011). *Haskell the Craft of Functional Programming*. Addison Wesley, third edition. ISBN 0201882957. URL http://www.haskellcraft.com/craft3e/Home.html. 28

van Rossum, Guido and Fred Drake (2011a). *An Introduction to Python*. Network Theory Limited, revised edition. ISBN 1906966133. 22

van Rossum, Guido and Fred Drake (2011b). *The Python Language Reference Manual*. Network Theory Limited, revised edition. ISBN 1906966141.

ToC

Python Code Index

Index for some (but not all) of the python code. Note that the index commands are placed after any *listings* environment containing the code to be indexed.

ToC

Diagrams Index

Index for some of the PGF/TikZ diagrams. Note that the indexing commands are placed after the diagram code

ToC